REPORTS OF PRACTICAL ONCOLOGY AND RADIOTHERAPY 22 (2017) 455–462

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: http://www.elsevier.com/locate/rpor

Review

Internet-based computer technology on radiotherapy

James C.L. Chow*

Department of Radiation Oncology, University of Toronto and Radiation Medicine Program, Princess Margaret Cancer Centre, University Health Network, Toronto, ON M5G 2M9, Canada

ARTICLE INFO

Article history:
Received 4 November 2016
Received in revised form
7 February 2017
Accepted 21 August 2017
Available online 8 September 2017

Keywords:
Radiotherapy
Computer technology
Cloud computing
Machine learning
Big data

ABSTRACT

Recent rapid development of Internet-based computer technologies has made possible many novel applications in radiation dose delivery. However, translational speed of applying these new technologies in radiotherapy could hardly catch up due to the complex commissioning process and quality assurance protocol. Implementing novel Internet-based technology in radiotherapy requires corresponding design of algorithm and infrastructure of the application, set up of related clinical policies, purchase and development of software and hardware, computer programming and debugging, and national to international collaboration. Although such implementation processes are time consuming, some recent computer advancements in the radiation dose delivery are still noticeable. In this review, we will present the background and concept of some recent Internet-based computer technologies such as cloud computing, big data processing and machine learning, followed by their potential applications in radiotherapy, such as treatment planning and dose delivery. We will also discuss the current progress of these applications and their impacts on radiotherapy. We will explore and evaluate the expected benefits and challenges in implementation as well.

1. Aim

The aim of this paper is to review some recent Internet-based computer technologies, such as cloud computing, big data processing and machine learning, and their potential applications in radiotherapy. The concept and recent progress of the above technologies will be illustrated and their implementations in radiotherapy, in particular treatment planning and dose delivery, will be explored. We will review some examples of application and their benefits to the radiation dose delivery. We will also review the general procedure and discuss the difficulty to implement those novel technologies in radiation cancer treatment. The impacts and challenges of using the above technologies in radiotherapy will be discussed.

2. Introduction

Computing devices are machines which can carry out logical operations automatically based on an input program or code. Because of this special ability, computers have been

* *Correspondence to*: Department of Radiation Physics, Princess Margaret Cancer Centre, 700 University Ave., Toronto, ON M5G 2M9, Canada.

E-mail address: james.chow@rmp.uhn.on.ca

http://dx.doi.org/10.1016/j.rpor.2017.08.005
1507-1367/

used world-wide as a control console in various industrial and health systems. The use of computers to replace humans for routine jobs can save human resources, increase production speed and reduce human errors. Computer hardware is making rapid advancement. The well-known Moore's law predicted that the number of transistors in a dense integrated circuit doubles about every two years.[1] Although nowadays this law has issues regarding some novel developments in mobile devices and data centers, the speed of computer technological advancement has been huge in recent years, resulting in many new applications in radiation dose delivery.

In radiotherapy, the typical treatment chain includes diagnosis and consultation, patient positioning and immobilization, computed tomography scanning or simulation, target and critical organ contouring and treatment planning, patient setup verification, and dose delivery. Computer is used in every component of the chain to perform hardware control, dose calculation, image processing/registration, data management and communication. Such radiation treatment chain is closely monitored by a quality assurance system, a program to control and maintain the standard of quality in the treatment to prevent error. With the increase of computing speed due to the innovation of parallel processing, dose calculation time in treatment planning is reduced.[2–4] This shortens patient treatment time and, thus, increases patient throughput. Increasing computer speed also makes real time image processing/deformation possible. This benefits the quality assurance in patient setup before daily treatment using various imaging modalities (e.g. computed tomography, magnetic resonance imaging and ultrasound) in image-guided radiotherapy.[5–7]

Some computer applications in radiotherapy, such as parallel processing and grid computing, have been making stable progress since they were applied to radiation dose calculation in treatment planning.[8–11] In this review, we will focus on some recent Internet-based computer technologies with abilities beyond general hardware control and computing. These technologies include cognitive computing and big data with innovations engendered by the recent rapid developments of information technology and artificial neural networks.

3. Cloud computing

In the past, when Internet was less developed and access to the Internet was less popular, individual users had to purchase their workstation hardware and software as per their need to carry out computer tasks. They not only had to spend money and time on upgrading the data storage and software, but also needed to bear the risk of computer system shutdown and response for the repairing and maintenance. Data sharing between users was limited at that time due to the lack of Internet infrastructure and low data transfer speed. Since the introduction and application of the TCP/IP protocol in networking, Internet became more popular and triggered the development of web-based application.[12] Cloud computing is a web-based technology where users share hardware and software in the cloud. Individual users only need to connect to the cloud, covering all complex network elements for simplicity, from their local workstation through the Internet, and they can perform different computer tasks according to their needs.[13,14] This cloud is offered by some service providers such as Amazon Web Services™, Google Compute Engine™, Windows Azure™ and Aruba Cloud™, which employ a "pay as you go" model. The users pay for running their computer task as per its size of database, required number of compute nodes (processors) and length of computing time. Using cloud computing, users do not need to take care of the related hardware and software update, repair and maintenance because the service provider will handle them. This "pay as you go" approach also ensures the flexibility for the users to manage their budget so that more effective resource allocation would be possible.

The cloud provides different software and platforms to support application, database for storage, compute nodes for calculation and system infrastructure development. Specific applications tailor-made by the user are possible using the infrastructure development platform provided by the cloud. Basically, cloud computing provides the following service models: software as a service (SaaS), platform as a service (PaaS) and Infrastructure as a service (IaaS). The SaaS model only provides the users with applications running on the cloud. The users can access the above applications through a web browser but they cannot manage the underlying infrastructure. Web-based email is an example. The PaaS model, however, offer a platform with programming languages, libraries and tools so that users can create and develop their own applications. However, they still cannot control the underlying infrastructure, including the operating system, network and storage. To control provision processing, data storage and computing resource to run arbitrary software, users need to use the IaaS model which offers the computing infrastructure. Examples of IaaS include Amazon EC2™, Windows Azure™ and Google Compute Engine™. The cost efficiency of using cloud computing in radiotherapy can be demonstrated by an example from proton treatment planning. In a proton therapy center, if each patient plan needs about 100 h of computing cluster time and there are 1,000 patients per year, the total CPU time is 100,000 h. Assuming a three-year life time of the computing cluster with cost equal to $1,000 per compute node plus $200 in maintenance per year per node, the total cost in three years is $160,000 for a cluster containing 100 compute nodes or $53,000 per year. Using cloud computing, for $0.1 per CPU-h with additional 5% of storage and data transfer, the total cost is only $10,500 per year. This is about 80% cheaper compared to the cost of $53,000 for regular computing hardware.[15]

3.1. *Cloud computing in radiotherapy*

There are studies on the applications of cloud computing technology on dose calculations in radiation treatment planning. This takes advantage of parallel processing in the cloud using flexible number of compute nodes. Wang et al. performed a Monte Carlo dose calculation using photon and electron beams on a cloud computing infrastructure.[16] They used the EGS5 Monte Carlo code[17] and proved that the output of cloud-based simulations is identical to the single-threaded implementation. They, therefore, concluded that cloud computing can provide high performance parallel processing in dose calculation for radiation treatment planning. On the other hand, Poole et al. described their implementation of cloud computing in dose calculation using the GEANT4 Monte Carlo code.[18] They examined the simulation completion time and found that it decreased as $1/n$, where n is the number of parallel machines. Poole et al. demonstrated that cloud computing can be used in rapid Monte Carlo dose calculation without the need for local computer hardware in radiotherapy. Chow carried out a study on implementing cloud computing in preclinical radiation treatment planning.[19] The infrastructure of the treatment planning system in the cloud using the Monte Carlo method as a dose calculation engine was explored, with a test of the dose computing time against the number of compute nodes. The EGSnrc Monte Carlo code[20] was used and it was found that the computing time does not increase proportionally to the number of compute nodes, but there exists an optimization point beyond which further increase in the number of compute nodes would not decrease the computing time effectively. Chow also examined the performance of 4D radiation treatment planning using Monte Carlo simulation on the cloud.[21] Chow evaluated the planning performance through optimizations of the number of compute nodes, number of computed tomography image sets and dose construction time in image voxel tracking. It is found that the dependence of computing time on the number of nodes was affected by the diminishing return of the number of nodes selected in Monte Carlo simulations. Furthermore, a web application called CloudMC running on the cloud platform was developed by Miras et al.[22] using the Window Azure™. Miras et al. tested the CloudMC using the Penelope Monte Carlo code[23] and found that the Monte Carlo computing time decreased with the number of instances following Amdahl's law.[24] They concluded cloud computing will play a significant role in radiation dose calculation using Monte Carlo simulation. For radiation treatment planning considering both the dose calculation and plan optimization, Na et al.[25] established a cloud computing infrastructure using Amazon EC2™. Their cloud computing environment can speed up in about 14-fold for dose kernel calculation and plan optimization in different cancer treatment sites such as head-and-neck, lung and prostate. They concluded that their web-based treatment planning system can substantially improve the inverse planning speed. Lai et al. introduced a new cloud service model called knowledge as a service (KaaS) that can be applied in radiotherapy.[26] They used a case study of medical service industry in China to demonstrate that their suggested KaaS model can provide the best practices through collaboration activated among the radiation treatment, molecular imaging and knowledge center as platforms to share and access information.

3.2. *Ethical and security issues in cloud computing for radiotherapy*

Since the cloud service providers have multiple users, there is a concern about the privacy and security of transferring and storing the patient data in the cloud through Internet. The central question is to make sure only authorized users can access patient data. Such data protection issue should be assured by the service provider based on a written contract. In addition, a cancer center or hospital providing patient data should have the right to maintain their ownership of the data. For example, the service provider has to destroy all patient data and related backup permanently when instructed by the data owner. There are laws regarding patient's privacy, such as the Health Insurance Portability and Accountability Act (HIPPA) in the USA, and the Personal Information Protection and Electronic Documents Act in Canada. The cloud service providers, such as Amazon, have set up guidelines to assist users for compliance with the American HIPPA regulations covering patient data privacy in the USA. On the other hand, from the technological point of view, more advanced Internet techniques, such as SSL and TLS, are used to provide data transfer and exchange.[27] This can increase the level of data protection. Moreover, different secure VPN protocols, such as PPTP, L2TP/IPSec, OpenVPN, SSTP and IKE/IKEv2, have been developed to provide a very high network security between the user and cloud.[28] For data storage, the redundant array of independent nodes (RAIN) implemented the redundant/reliable array of inexpensive/independent disks (RAID) that can provide fully automated data recovery in the network.[29] This technology can help the cloud service provider to maintain data security.

4. Big data processing

The rapid advancement of data technology in recent years has led to a marked decrease of cost and substantial increase of storage capacity. The bandwidth, reflecting the cost per megabyte of data storage, is getting larger over time. Still, the ever increasing requirement of data volume due to the usage of smartphone and tablet challenges the developments of data capacity and processing. The innovation of cloud computing moves the data storage from the user's local workstation to the remote cloud maintained by the service provider. This flexible and new approach based on Internet makes data transfer, storage and processing very convenient to acquire, and all data can be accumulated from a general task. The name big data is therefore used to describe a dataset that is too large and complex to be dealt with by a traditional data processing system.[30] The complexity of big data can be visualized by a 3 "V" model describing the data growth challenge. The 3 "V" represents the rapid growth of volume (amount of data), velocity (data transfer speed) and variety (data type and sources) of data.[31]

In radiotherapy, we can find from the patients' database treatment planning and dose delivery information, such as radiation beam settings, dose prescriptions, dose distributions and dose delivery records. The Electronic Medical Records is an example of such a patient treatment database.[32] We can also find patients' image set of different diagnostic and therapeutic imaging modalities from image databases, such as the Picture Archiving and Communication System.[33] However, the above clinical data can become even bigger by including genomics, proteomics and metabolomics information as well as various tissue and blood testing results used in chemotherapy. Apart from the treatment data in radiotherapy, chemotherapy and surgery, big data in cancer treatment also contain all information from clinical consultation. This includes data from pathology, diagnosis, patients' performance statues and comorbidities. In addition, big data should have patients' records of follow up including clinical assessments, disease statues, qualities of life and survivals. Application of big data in radiotherapy would involve data from other treatment modalities, such as chemotherapy and surgery. Therefore, by considering all aspects of patients in big data, better treatment outcome can be obtained based on the improved clinical decision.

4.1. *Application of big data in radiotherapy*

The application of big data in radiotherapy has many advantages. Deng explored the challenges and opportunities of using big data in radiotherapy.[34] He explained the essential composition of big data and how it can be applied in radiotherapy with examples. These included a population-based study to identify the best strategy in prostate cancer treatment and the development of a multi-institutional standardized database for radiotherapy quality assurance. He concluded that with the continuous progresses of effective and appropriate tools mastering the big data in radiation oncology, better clinical decision making, and personalize treatment strategy should be achieved. The improved patient safety and treatment efficacy due to the application of big data can result in a more cost-effective healthcare model. Although he predicted the complexity of the model would gradually be increased overtime, the exchange process could lead to a rapid knowledge generation in radiotherapy. The applications of big data in radiation oncology for research, quality assessment and clinical care were examined by Benedict et al.[35] They pointed out that the traditional multi-institutional clinical trial, which accounts only for about 3% of all patient data, can greatly be enhanced by using big data, including the rest of 97% of so called "dark data". However, releasing this invisible data requires a collaboration to enforce a standardized lexicons and proper curating in routine data mining. Further challenges in implementing big data in radiotherapy were explored by Benedict et al. focusing on various data types from multiple cancer centers/hospitals used in research, quality assessment and clinical care.[36] They suggested using big data to formulate a knowledge health management system that can help and improve patient treatment. Rosenstein et al. illustrated how big data application could improve the clinical and basic research in radiotherapy.[37] They discussed different radiation oncology problems where benefits could be achieved by considering big data. These included cancer sites such as the head- and-neck, lung, breast, prostate and rectum. They gave an example of cloud-based informatics infrastructure for data transfer, quality assurance evaluation and data integration, and introduced the knowledge-guided radiotherapy, which applies big data to predict better patient outcome. Coates et al. focused on big data analytics for prostate radiotherapy.[38] They discussed the data mining of heterogeneous types, such as the patient-specific clinical parameters, treatment-related dose-volume metrics and radiobiological risk factors. Using newly developed computer technologies, such as machine learning and artificial intelligence, they predicted data analytics would take an important role in prostate radiotherapy in the future. Chow proposed a hybrid big data cloud to facilitate radiation treatment plan evaluation and comparison.[39] A web-based interface was designed to link to the cloud containing different treatment sites such as the head-and-neck, lung, breast and prostate. Chow concluded that such an on-growing dynamic data cloud can help radiotherapy planners in a more accurate plan evaluation.

4.2. *Challenges in application of big data in radiotherapy*

Since application of big data in radiotherapy generally requires data sharing from multiple cancer centers/hospitals, there are both technical and business concerns about data adaptability, ownership, security, privacy and publishing right. Huser et al. studied the challenges in big data for radiotherapy and highlighted issues of data volume, institutional review board, semantic data integration, databases versus data files and access to knowledge bases.[40] From their investigations on current approaches to those challenges, they found that newly trained data scientists are shifting from data acquisition to transformation due to the emergence of the data wrangling concept. Skripcak et al., on the other hand, illustrated a big data framework for international research data exchange in radiotherapy and oncology.[41] They pointed out that data interoperability among multiple cancer centers/hospitals is a major concern. They further explored the challenges of utilization of standards, data quality and privacy concerns, data ownership, rights to publish, data pooling architecture and storage in establishing the data pooling model and data exchange strategy. They concluded that creating a specific data exchange strategy among cancer centers/hospitals is challenging but feasible. Potters et al. showed how to build up a radiation oncology learning system.[42] They found that the construction of a system should include a change of philosophy that quality assurance should use a systematic approach, instead of replying only on an expert panel. The support from related hardware vendors and software manufacturers to follow a new standard for the database is also needed.

5. Machine learning

In radiotherapy, treatment planning involves image set of patient's anatomy for radiation beam setup and dose calculation. There are different image modalities such as computed tomography, magnetic resonance imaging, single-photon emission computed tomography and positron emission tomography. Image processing techniques, such as registration and deformation, are developing fast to integrate data from the above image modalities together so that most updated and detailed information can be produced. In this process, it would be great to have automatic recognition of patient's organs in the treatment system in planning, and useful to have computing technique such as pattern recognition.[43] However, pattern recognition depends on experts to program a smart algorithm which can solve the problem. Such an algorithm can be built up by machine learning using mined data.[44] The machine is guided (supervised) or not guided (unsupervised) by the user through continuously feeding new data to a machine learning algorithm and waiting for the computations to finish. This machine learning method takes advantage of big data cloud advancement which makes big data mining and processing easy. Deep machine learning or deep learning is a new branch in machine learning using a set of algorithms based on complex neural networks.[45] The word "deep" reflects the number of layers of the neural networks used in the learning process. By using some powerful networks, such as convolutional multi-layer neural network,[46] large scale image recognition task becomes possible through a series of convolutions, pooling and classification. Although the scale of data is big and the neutral network containing various formulas is complex, the availability of cloud computing makes this novel computer technology possible.

Deep learning used a multiple complex linear map layers to perform feature learning through high level abstraction modeling from data. As mentioned above, in deep learning, the handcrafted features with efficient algorithms can be replaced by semi-supervised or supervised feature learning so that the machine can learn from the observation automatically. The observation can be represented by different ways. For an image, the observation can be converted to vector of intensity values per pixel, or different sets of edges and shape. This data will be input into learning architectures constructed by layers of neural networks composed of neuron representing linear functions. There are many well developed networks, such as deep neural networks, convolutional deep neural networks and recurrent neural networks.[47] They have their own characteristics in feature learning. For example, convolution deep neural networks are popular to be used in graphical deep learning because it is flexible to handle various image features in different regions of the image.

The learning process is a continuous linear map optimization in neural networks aiming at reducing the deviation between the computing result and the observation target. The computing continuously searches for the global minimum representing the minimum deviation between the result and target. It can be seen that a lot of observation data can be input into the machine so as to increase its knowledge to the target.

5.1. *Application of machine learning in radiotherapy*

Machine learning was used in treatment planning and studied by different groups. Petrovic et al. investigated the knowledge-light adaptation approaches using their suggested case-based reasoning system in treatment planning.[48] They tested approaches based on machine learning and adaptation-guided retrieval using brain cancer cases treated with the 3D-conformal technique. They found that neural networks-based adaptation improved the success rate of the case-based system by 12%, while the adaptation-guided retrieval of the case for beam number improved the success rate of the system by 29%. Based on their results of treatment plans, they concluded their proposed adaptation methods can improve the performance of their case-based system regarding the number of beams but beam angles. Stanhope et al. presented a knowledge-based technique to find out the degradations in dose-volume histogram (DVH) for prostate radiotherapy.[49] In their work, machine learning of 198 previously treated patients' plans were adapted to test patient's anatomy to determine the patient-specific DVH ranges using the single-arc volumetric modulated arc therapy technique. Plan comparison was carried out among the original DVH, degraded DVH due to delivery error and predicted DVH from machine learning. Changes in dosimetric indices from DVH in terms of figures of merit (FOM) as dosimetric indices were determined for plan evaluation. They found that their knowledge-based quality assurance approach allowed customized quality assurance criteria specified for treatment site and institution. Such approach was more sensitive to errors than the quality assurance criteria traditionally based on organ complication rates. On the other hand, Chow et al. suggested using a parameter called dose-volume consistency to estimate the dose-volume variability of DVH curves in the treatment plan database.[50] They examined the plan quality consistency between the prostate patients using the intensity modulated radiotherapy and volumetric modulated arc therapy technique. From the patient data, they found that a small prostate tumor control probability variation can be maintained by reducing the dose-volume variability among the patients. They also concluded that dependences of the rectal equivalent uniform dose and rectal normal tissue complication probability on the dose-volume consistency were not significant for both delivery techniques. Using machine learning, Dean et al. constructed a predictive model of mucositis in head-and-neck radiotherapy.[50,51] They used the support vector and random forest classification in the model and calculated the normal tissue complication probability and spatial dose metrics in the treatment. They concluded that reducing the irradiated volume of the oral cavity may reduce the chance to have mucositis. Guidi et al. used machine learning to predict patients who would benefit from adaptive radiotherapy and replanning intervention.[52] Ninety head-and-neck cancer patients were included in a multicenter dataset. They developed a machine-learning classifier to analyze dose variations of parotid glands and used the support vector machines for the time-series evaluation. They found that replanning would be required after the 4th week based on the parotid gland tolerance. Guidi et al. concluded that their decision-making tool can be applied to overcome adaptive radiotherapy challenges, because using their tool enables ideal time for replanning intervention to be predicted. Apart from machine learning application in treatment planning, Carlson et al. applied machine learning in prediction of the multi-leaf collimator positional errors.[53] They trained the machine by using the delivered multi-leaf collimator log data. Through comparison with the planned collimator data, the machine can predict the positional error. Carlson et al. found that including such positional error in the patient specific quality assurance can increase the accuracy of gamma test by 4.17%. They concluded that with the help from machine learning, treatment planner can obtain a more realistic view of dose distribution by considering the multi-leaf collimator positional error in dose delivery.

5.2. *Prospective issues of machine learning in radiotherapy*

Although there are some works on the application of machine learning in radiotherapy, the adoption speed is still slow. This is due to the lack of understanding in the background, resource, and technique in the research practice. It should be noted that machine learning requires training on a machine using big data in the cloud. As those Internet computer technologies are still new to the radiation oncology researchers who used to conduct traditional clinical trials, studying such rapidly developing technologies would take time. It will also take time to select a suitable machine learning method specific to solving a problem. Experience and skills are therefore slowly to build up through repeated comparison and validation using other well-known models.[54] However, it is seen that some popular machine learning models, such as support vector machine and artificial neural network, have been used to predict radiotherapy outcome.[55] This shows that the machine learning technology is getting matured.

Another issue for the application of machine learning in radiotherapy is the complex data-mining process. For example, calculation of the normal tissue complication probability and tumor control probability can be enhanced by not only considering the dosimetric but non-dosimetric data. This non-clinical data include variables of age, sex and histology, which are new for the researcher.[56] In addition, data mining of the above information needs to consider the format, privacy, ownership and security of the source and it is complicated. It is desired to have collaboration between multiple cancer centers and manufacturers to standardize the data collection process so as to simplify the complexity of data-mining.[57]

6. Legal aspects and proposals

Legal aspect, medical software engineering labeling and validation process before the clinical use of Internet-based service are challenging to health workers. In radiotherapy, this is particularly important when certain sub-tasks are outsourced from the source institution to the service provider, and legal accountability situation became difficult as none of the parties has full control of the entire environment. The security/legal issues can be classified into three categories, namely: traditional security concerns (e.g. network intrusion and attack), availability issues (service center downtime and outage) and third party data control-related issues (legal implications of data and applications being held by a third party). Solutions for the security issues can be focused on some novel security approaches that allow the service provider to have some control over the user's data. The proposals included information-centric security to allow data to be self-describing and defending, privacy-enhanced intelligence to further encrypt the data, and implementation of trusted computing to offer a statement of compliance produced by the monitor.[58]

7. Future directions

Cloud computing can provide web-based services to data storage, application and calculation. It can also provide a platform for the user to develop their infrastructure and application. As more and more radiotherapy researchers recognized the advantages of cloud computing, different studies have been carried out and reviewed here. These included dose calculations employing the parallel processing technique, big data science and knowledge-based quality assurance system development. The innovation rate of cloud computing has been very high recently. It can be seen that cloud computing started to abstract infrastructure and operation processes from physical to virtual. The innovation of deep machine learning can convert our research process from analytics to machine learning using big data. This simplifies the research process so that results can be acquired faster and be more accurate. In the future, as cloud computing becomes more user friendly, researchers without a lot of knowledge on the complex Internet-based technologies, such as machine learning, can work easily on the cloud to carry out studies based on big data. This will require an excellent user interface.

With the rapid development of data-mining technique and capacity of the cloud, big data processing becomes popular. In radiotherapy, big data not only includes dosimetric information, such as the dose distribution, beam setup and dose prescription, but also non-dosimetric data from other treatment modalities, such as surgery and chemotherapy. In addition, it can include data from patient's consultation and follow-up. This means that data volume will continue to grow in the future with accumulation of more patient cases, and studies on the data allocation and storage will become significant and urgent. In fact, it is found that some studies may not need to access all data and perhaps the concept of fast and actionable data will be focused on.[59] To process the data, tools for data analyze should be improved. User-friendly tool based on graphical user interface is expected to help users who have no computer coding experience to build up their data processing applications.[60] Moreover, it should be studied how to effectively use big data. This involves the data filtering technique in data-mining.[61] Some other future works like data privacy and security, real-time decision making and virtual machine training will continue to develop.

Machine learning has started to analyze patient data in radiotherapy and proved to be effective compared to the traditional approach. However, results and decisions made by the machine may not be understood by the researcher due to a very complex artificial neural network and algorithm.[62] This may bother the radiotherapy studies because researchers cannot explain the results based on a model or formula. It is highly expected in the future that users can easily and comprehensively access the machine to find out the related algorithm for the final result. Although there are currently some powerful algorithms, such as elastic-net, support-vector machine and neural network in machine learning, more tailor-made algorithms which can be fitted to the radiotherapy patient data structure should be developed. It can be seen that development of a learning mechanism, such as increasing the number of layers in the neural network in deep machine learning makes the system more complex. However, it is expected that the user who does not have knowledge on the underlying structure of the learning algorithm, can master the virtual machine easily to produce a useful result for the advancement of radiation oncology study.[63]

8. Conclusion

Recently, Internet-based computer technology such as cloud computing, big data processing and machine learning have

Table 1 – Summary of possible outsourced tasks using the internet-based computing technologies in radiotherapy.

Possible outsourced tasks	Main internet-based techniques	References
External beam treatment planning – dose calculation and plan optimization	Cloud computing	8,9,15,16,18,19,21,22,25
Radiation treatment system and delivery	Cloud computing, big data	26,35,36
Clinical study based on electronic medical record and multi-institutional data sharing	Cloud computing, big data, machine learning	32,34–39,41,54–57,61,62
Safety and quality improvement in radiation oncology	Big data	35,36,42,51
Pattern recognition in medical imaging	Machine learning, big data	46
Treatment planning quality assurance	Machine learning, big data	48–50,52
Multi-leaf collimator positional accuracy evaluation	Machine learning	53

been used in radiotherapy study. Although such applications are still not popular, some preliminary studies have been carried out and reviewed here. A summary of the possible tasks outsourced to the Internet-based services with related references is shown in Table 1. Application of these technologies, for example, machine learning trained by big data in the cloud, leads to a knowledge-based system which can make the radiation dose delivery process automatic. There are challenges to apply these technologies in radiotherapy, such as lack of personnel who knows both Internet-based technologies and radiotherapy. Other challenges include collaboration among multiple institutions for data privacy, security and ownership, support from the vendors and manufacturers for standardization of the dataset and tools, and verification of the results. There is a concern that much longer time would need to be spent on solving the above problems, which makes it challenging to keep pace with the advancement of the technologies. However, it has been proved that a knowledge-based system can benefit the radiotherapy process. Therefore, implementation of the Internet-based technologies in a cancer center should start sooner rather than later to keep pace with such rapid technological progress.

Conflict of interest

None declared.

Financial disclosure

None declared.

REFERENCES

1. Moore GE. Progress in digital integrated electronics. In: *Technical digest, IEEE international electron devices meeting*. 1975. p. 11–3.
2. Jia X, Ziegenhein P, Jiang SB. GPU-based high-performance computing for radiation therapy. *Phys Med Biol* 2014;**59**:R151–82.
3. Gu X, Jelen U, Li J, Jia X, Jiang SB. GPU-based pencil beam algorithm with 3D-density correction for radiotherapy dose calculation. *Phys Med Biol* 2011;**56**:3337–50.
4. Chow JCL. A performance evaluation on Monte Carlo simulation for radiation dosimetry using cell processor. *J Comput Methods Sci Eng* 2011;**11**:1–12.
5. Arabloo J, Hamouadeh P, Mousavinezhad SM, et al. Health technology assessment of image-guided radiotherapy (IGRT): a systematic review of current evidence. *Med J Islam Repub Iran* 2016;**30**:318.
6. Sterzing F, Engenhart-Cabillic R, Flentje M, Debus J. Image-guided radiotherapy: a new dimension in radiation oncology. *Dtsch Arztebl Int* 2011;**108**:274–80.
7. Jaffray D, Kupelian P, Djemil T, Macklis RM. Review of image-guided radiation therapy. *Expert Re Anticancer Ther* 2007;**7**:89–103.
8. Vadapalli R, Yepes P, Newhauser W, Lichti R. Grid-enabled treatment planning for proton therapy using Monte Carlo simulations. *Nucl Technol* 2011;**175**:16–21.
9. Downes P, Yaikhom G, Giddy JP, et al. High-performance computing for Monte Carlo radiotherapy calculation. *Philos Trans A Math Phys Eng Sci* 2009;**367**:2607–17.
10. Kalantzis G, Tachibana H. Accelerated event-by-event Monte Carlo microdosimetric calculations of electrons and protons tracks on a multi-core CPU and a CUDA-enabled GPU. *Comput Methods Programs Biomed* 2014;**113**:116–25.
11. Souris K, Lee JA, Sterpin E. Fast multipurpose Monte Carlo simulation for proton therapy using multi- and many-core CPU architectures. *Med Phys* 2016;**43**:1700.
12. Balakrishnan H, Stemm M, Seshan S, Padmanabhan V, Katz RH. TCP behavior of a busy internet server: analysis and solutions. In: *Proceedings of IEEE INFOCOMM'98*. 1998. p. 252–62.
13. Dikaiakos MD, Katsaros D, Mehra P, Pallis G, Vakali A. Cloud computing: distributed Internet computing for IT and scientific research. *IEEE Internet Comput* 2009;**13**:10–3.
14. Venters W, Whitley EA. A critical review of cloud computing: researching desires and realities. *J Inform Technol* 2012;**27**:179–97.
15. Keyes RW, Romano C, Arnold D, Luan S. *Radiation therapy calculations using an on-demand virtual cluster via cloud computing*; 2010, ar**Xiv preprint ar**Xiv:1009.5282.
16. Wang H, Ma Y, Pratx G, Xing L. Toward real-time Monte Carlo simulation using a commercial cloud computing infrastructure. *Phys Med Biol* 2011;**56**:N175–81.
17. Namito Y, Hirayama H. EGS5 code, outline and how to start using the code. *Igaku Butsuri* 2013;**33**:82–7.
18. Poole CM, Cornelius I, Trapp JV, Langton CM. Radiotherapy Monte Carlo simulation using cloud computing technology. *Australas Phys Eng Sci Med* 2012;**35**:497–502.
19. Chow JCL. Application of cloud computing in preclinical treatment planning. In: Clary TS, editor. *Horizons in computer science research volume 10*. New York: Nova; 2015. p. 35–48.
20. Kawrakow I, Mainegra-Hing E, Rogers DWO, Tessier F, Walters BRB. *The EGSnrc code system: Monte Carlo simulation of electron and photon transport*; 2016. NRCC Report PIRS-701.
21. Chow JCL. Performance optimization in 4D radiation treatment planning using Monte Carlo simulation on the cloud. *J Comput Methods Sci Eng* 2016;**16**:147–56.

22. Miras H, Jimenez R, Miras C, Goma C. CloudMC: a cloud computing application for Monte Carlo simulation. *Phys Med Biol* 2013;**58**:N125–33.
23. Baro J, Sempau J, Fernandez-Varea JM. PENELOPE: an algorithm for Monte Carlo simulation of the penetration and energy loss of electrons and positrons in mater. *Nucl Instrum Methods Phys B* 1995;**100**:31–46.
24. Hill MD, Marty MR. Amdahl's law in the multicore era. *Computer* 2008;**41**:33–8.
25. Na YH, Suh TS, Kapp DS, Xing L. Toward a web-based real-time radiation treatment planning system in a cloud computing environment. *Phys Med Biol* 2013;**58**:6525–40.
26. Lai IKW, Tam SKT, Chan MFS. Knowledge cloud system for network collaboration: a case study in medical service industry in China. *Expert Syst Appl* 2012;**39**:12205–12.
27. Berbecaru D, Lioy A. *On the robustness of applications based on the SSL and TLS security protocols*. Berlin, Heidelberg: Springer-Verlag; 2007.
28. Sanzgiri K, Dahill B, Levine BN, Shields C, Belding-Royer EM. A secure routing protocol for ad hoc networks. In: *Network protocols proceedings. 10th IEEE international conference on 2002 November 12*. IEEE; 2002. p. 78–87.
29. Fan B, Tantisiriroj W, Xiao L, Gibson G. DiskReduce: RAID for data-intensive scalable computing. In: *Proceedings of the 4th annual workshop on petascale data storage*. 2009. p. 6–10.
30. De Mauro A, Greco M, Grimaldi M. What is big data? A consensual definition and a review of key research topics. In: *AIP conference proceedings*. 2015. p. 97–104, vol. 1644, No. 1.
31. Beyer M. Gartner says solving 'Big Data' challenge involves more than just managing volumes of data. *Gartner* 2011;(July).
32. Dean BB, Lam J, Natoli JL, Butler Q, Aguilar D. Use of electronic medical records for health outcomes research. *Med Care Res Rev* 2009;**66**:611–38.
33. Arenson RL. Picture archiving and communication systems. *West J Med* 1992;**156**:298–9.
34. Deng J. Big data in radiation oncology: challenges and opportunities. *Cancer Sci Res Open Access* 2014;**1**:1–2.
35. Benedict SH, El Naga I, Klein EE. Introduction to big data in radiation oncology: exploring opportunities for research, quality assessment, and clinical care. *Int J Radiat Oncol Biol Phys* 2016;**95**:871–2.
36. Benedict SH, Hoffman K, Martel MK. Overview of the American Society for Radiation Oncology-National Institutes of Health-American Association of Physicists in Medicine Workshop 2015: exploring opportunities for radiation oncology in the era of big data. *Int J Radiat Oncol Biol Phys* 2016;**95**:873–9.
37. Rosenstein BS, Capala J, Efstathiou JA, et al. How will big data improve clinical and basic research in radiation therapy? *Int J Radiat Oncol Biol Phys* 2016;**95**:895–904.
38. Coates J, Souhami L, El Naga I. Big data analytics for prostate radiotherapy. *Front Oncol* 2016;**6**:149.
39. Chow JCL. Evaluation of radiation treatment plan using big data and cloud computing. In: *Conference reports from diTHINK 2016*. 2016.
40. Huser V, Cimino JJ. Impending challenges for the use of big data. *Int J Radiat Oncol Biol Phys* 2016;**95**:890–4.
41. Skripcak T, Belka C, Bosch W, et al. Creating a data exchange strategy for radiotherapy research: towards federated databases and anonymized public datasets. *Radiother Oncol* 2014;**113**:303–9.
42. Potters L, Ford E, Evans S, Pawlicki T, Mutic S. A system approach using big data to improve safety and quality in radiation oncology. *Int J Radiat Oncol Biol Phys* 2016;**95**:885–9.
43. Shirato H, Shimizu S, Kitamura K. Four-dimensional treatment planning and fluoroscopic real-time tumour tracking radiotherapy for moving tumor. *Int J Radiat Oncol Biol Phys* 2000;**48**:435–42.
44. Michalski RS, Carbonell JG, Mitchell TM, editors. *Machine learning: an artificial intelligence approach*. Springer Science & Business Media; 2013.
45. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;**521**:436–44.
46. Oquab M, Bottou L, Laptev I, Sivic J. Learning and transferring mid-level image representations using convolutional neural networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014. p. 1717–24.
47. Schmidhuber J. Deep learning in neural networks: an overview. *Neural Netw* 2015;**61**:85–117.
48. Petrovic S, Khussainova G, Jagannathan R. Knowledge-light adaptation approaches in case-based reasoning for radiotherapy treatment planning. *Artif Intell Med* 2016;**68**:17–28.
49. Stanhope C, Wu QJ, Yuan L, et al. Utilizing knowledge from prior plans in the evaluation of quality assurance. *Phys Med Biol* 2015;**60**:4873–91.
50. Chow JCL, Jiang R, Kiciak A. Dose-volume consistency and radiobiological characterization between prostate IMRT and VMAT plans. *Int J Can Ther Oncol* 2016;**4**:447.
51. Chow JCL. Radiation treatment planning based on big data of previously treated plans using the Gaussian error function model. In: *Proceedings, HPCS 2015, advanced computing and big data: driving competitiveness and discovery*. 2015. p. 12.
52. Guidi G, Maffei N, Meduri B. A machine learning tool for re-planning and adaptive RT: a multicenter cohort investigation. *Phys Med* 2016. S1120-1797(16)30945-0.
53. Carlson JN, Park JM, Park SY, Park JI, Choi Y, Ye SJ. A machine learning approach to the accurate prediction of multi-leaf collimator positional errors. *Phys Med Biol* 2016;**61**:2514–31.
54. Yahya N, Ebert MA, Bulsara M, et al. Statistical-learning strategies generate only modestly performing predictive models for urinary symptoms following external beam radiotherapy of the prostate: a comparison of conventional and machine-learning methods. *Med Phys* 2016;**43**:2040.
55. Çınar M, Engin M, Engin EZ, Ateşçi YZ. Early prostate cancer diagnosis by using artificial neural networks and support vector machines. *Expert Syst Appl* 2009;**36**:6357–61.
56. Dean JA, Wong KH, Welsh LC, et al. Normal tissue complication probability (NTCP) modelling using spatial dose metrics and machine learning methods for severe acute oral mucositis resulting from head and neck radiotherapy. *Radiother Oncol* 2016;**120**:21–7.
57. Meldolesi E, van Soest J, Damiani A, et al. Standardized data collection to build prediction models in oncology: a prototype for rectal cancer. *Future Oncol* 2016;**12**:119–36.
58. Sen J. Security and privacy issues in cloud computing. In: *Architectures and protocols for secure information technology infrastructures*.; 2013. p. 1–45.
59. Kaisler S, Armour F, Espinosa JA, Money W. Big data: issues and challenges moving forward. In: *46th Hawaii international conference on system sciences (HICSS), 2013*. IEEE, January; 2013. p. 995–1004.
60. Chow JCL. Some computer graphical user interfaces in radiation therapy. *World J Radiol* 2016;**8**:255–67.
61. Liao SH, Chu PH, Hsiao PY. Data mining techniques and applications – a decade review from 2000 to 2011. *Expert Syst Appl* 2012;**39**:11303–11.
62. Kang J, Schwartz R, Flickinger J, Beriwal S. Machine learning approaches for predicting radiation therapy outcomes: a clinican's perspective. *Int J Radiat Oncol Biol Phys* 2015;**93**:1127–35.
63. Perez CA, Mutic S. Advances and future of radiation oncology. *Rep Pract Oncol Radiother* 2013;**18**:329–32.